

OIFE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/828,423

DATE: 11/26/2001
TIME: 15:04:21

Input Set : N:\Crf3\RULE60\09828423.txt
Output Set: N:\CRF3\11262001\I828423.raw

SEQUENCE LISTING

C--> 4 (1) GENERAL INFORMATION:

(i) APPLICANT: Hillman, Jennifer L.
Guegler, Karl J.
Patterson, Chandra

C--> 10 (ii) TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE
INHIBITOR HEAVY CHAIN PRECURSOR

11 (iii) NUMBER OF SEQUENCES: 5

13 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
16 (B) STREET: 3174 Porter Drive
17 (C) CITY: Palo Alto
18 (D) STATE: CA
19 (E) COUNTRY: USA
20 (F) ZIP: 94304

21 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette
24 (B) COMPUTER: IBM Compatible
25 (C) OPERATING SYSTEM: DOS
26 (D) SOFTWARE: Word Perfect 6.1/MS-DOS 6.2

27 (vi) CURRENT APPLICATION DATA:

29 (A) APPLICATION NUMBER: US/09/828,423
C--> 30 (B) FILING DATE: 05-Apr-2001
C--> 31 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 09/388,774
35 (B) FILING DATE:

36 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Cerrone, Michael C
39 (B) REGISTRATION NUMBER: 39,132
40 (C) REFERENCE/DOCKET NUMBER: PF-0505 US

41 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 650-855-0555
44 (B) TELEFAX: 650-845-4166
45 (C) TELEX:

46 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 942 amino acids
52 (B) TYPE: amino acid
53 (C) STRANDEDNESS: single
54 (D) TOPOLOGY: linear

55 (vii) IMMEDIATE SOURCE:

57 (A) LIBRARY: UTRSNOT02
58 (B) CLONE: 688183

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1 :

61 Met Leu Leu Leu Leu Gly Leu Cys Leu Gly Leu Ser Leu Cys Val
63 5 10 15
64

ENTERED

11/26/01

RAW SEQUENCE LISTING

DATE: 11/26/2001

PATENT APPLICATION: US/09/828,423

TIME: 15:04:21

Input Set : N:\Crf3\RULE60\09828423.txt

Output Set: N:\CRF3\11262001\I828423.raw

```

65 Gly Ser Gln Glu Glu Ala Gln Ser Trp Gly His Ser Ser Glu Gln
66                20                25                30
67 Asp Gly Leu Arg Val Pro Arg Gln Val Arg Leu Leu Gln Arg Leu
68                35                40                45
69 Lys Thr Lys Pro Leu Met Thr Glu Phe Ser Val Lys Ser Thr Ile
70                50                55                60
71 Ile Ser Arg Tyr Ala Phe Thr Thr Val Ser Cys Arg Met Leu Asn
72                65                70                75
73 Arg Ala Ser Glu Asp Gln Asp Ile Glu Phe Gln Met Gln Ile Pro
74                80                85                90
75 Ala Ala Ala Phe Ile Thr Asn Phe Thr Met Leu Ile Gly Asp Lys
76                95                100               105
77 Val Tyr Gln Gly Glu Ile Thr Glu Arg Glu Lys Lys Ser Gly Asp
78                110               115               120
79 Arg Val Lys Glu Lys Arg Asn Lys Thr Thr Glu Glu Asn Gly Glu
80                125               130               135
81 Lys Gly Thr Glu Ile Phe Arg Ala Ser Ala Val Ile Pro Ser Lys
82                140               145               150
83 Asp Lys Ala Ala Phe Phe Leu Ser Tyr Glu Glu Leu Leu Gln Arg
84                155               160               165
85 Arg Leu Gly Lys Tyr Glu His Ser Ile Ser Val Arg Pro Gln Gln
86                170               175               180
87 Leu Ser Gly Arg Leu Ser Val Asp Val Asn Ile Leu Glu Ser Ala
88                185               190               195
89 Gly Ile Ala Ser Leu Glu Val Leu Pro Leu His Asn Ser Arg Gln
90                200               205               210
91 Arg Gly Ser Gly Arg Gly Glu Asp Asp Ser Gly Pro Pro Pro Ser
92                215               220               225
93 Thr Val Ile Asn Gln Asn Glu Thr Phe Ala Asn Ile Ile Phe Lys
94                230               235               240
95 Pro Thr Val Val Gln Gln Ala Arg Ile Ala Gln Asn Gly Ile Leu
96                245               250               255
97 Gly Asp Phe Ile Ile Arg Tyr Asp Val Asn Arg Glu Gln Ser Ile
98                260               265               270
99 Gly Asp Ile Gln Val Leu Asn Gly Tyr Phe Val His Tyr Phe Ala
100               275               280               285
101 Pro Lys Asp Leu Pro Pro Leu Pro Lys Asn Val Val Phe Val Leu
102               290               295               300
103 Asp Ser Ser Ala Ser Met Val Gly Thr Lys Leu Arg Gln Thr Lys
104               305               310               315
105 Asp Ala Leu Phe Thr Ile Leu His Asp Leu Arg Pro Gln Asp Arg
106               320               325               330
107 Phe Ser Ile Ile Gly Phe Ser Asn Arg Ile Lys Val Trp Lys Asp
108               335               340               345
109 His Leu Ile Ser Val Thr Pro Asp Ser Ile Arg Asp Gly Lys Val
110               350               355               360
111 Tyr Ile His His Met Ser Pro Thr Gly Gly Thr Asp Ile Asn Gly
112               365               370               375
113 Ala Leu Gln Arg Ala Ile Arg Leu Leu Asn Lys Tyr Val Ala His

```

RAW SEQUENCE LISTING

DATE: 11/26/2001

PATENT APPLICATION: US/09/828,423

TIME: 15:04:21

Input Set : N:\Crf3\RULE60\09828423.txt

Output Set: N:\CRF3\11262001\I828423.raw

114		380		385		390
115	Ser Gly Ile Gly Asp Arg Ser Val Ser Leu Ile Val Phe Leu Thr					
116		395		400		405
117	Asp Gly Lys Pro Thr Val Gly Glu Thr His Thr Leu Lys Ile Leu					
118		410		415		420
119	Asn Asn Thr Arg Glu Ala Ala Arg Gly Gln Val Cys Ile Phe Thr					
120		425		430		435
121	Ile Gly Ile Gly Asn Asp Val Asp Phe Arg Leu Leu Glu Lys Leu					
122		440		445		450
123	Ser Leu Glu Asn Cys Gly Leu Thr Arg Arg Val His Glu Glu Glu					
124		455		460		465
125	Asp Ala Gly Ser Gln Leu Ile Gly Phe Tyr Asp Glu Ile Arg Thr					
126		470		475		480
127	Pro Leu Leu Ser Asp Ile Arg Ile Asp Tyr Pro Pro Ser Ser Val					
128		485		490		495
129	Val Gln Ala Thr Lys Thr Leu Phe Pro Asn Tyr Phe Asn Gly Ser					
130		500		505		510
131	Glu Ile Ile Ile Ala Gly Lys Leu Val Asp Arg Lys Leu Asp His					
132		515		520		525
133	Leu His Val Glu Val Thr Ala Ser Asn Ser Lys Lys Phe Ile Ile					
134		530		535		540
135	Leu Lys Thr Asp Val Pro Val Arg Pro Gln Lys Ala Gly Lys Asp					
136		545		550		555
137	Val Thr Gly Ser Pro Arg Pro Gly Gly Asp Gly Glu Gly Asp Thr					
138		560		565		570
139	Asn His Ile Glu Arg Leu Trp Ser Tyr Leu Thr Thr Lys Glu Leu					
140		575		580		585
141	Leu Ser Ser Trp Leu Gln Ser Asp Asp Glu Pro Glu Lys Glu Arg					
142		590		595		600
143	Leu Arg Gln Arg Ala Gln Ala Leu Ala Val Ser Tyr Arg Phe Leu					
144		605		610		615
145	Thr Pro Phe Thr Ser Met Lys Leu Arg Gly Pro Val Pro Arg Met					
146		620		625		630
147	Asp Gly Leu Glu Glu Ala His Gly Met Ser Ala Ala Met Gly Pro					
148		635		640		645
149	Glu Pro Val Val Gln Ser Val Arg Gly Ala Gly Thr Gln Pro Gly					
150		650		655		660
151	Pro Leu Leu Lys Lys Pro Tyr Gln Pro Arg Ile Lys Ile Ser Lys					
152		665		670		675
153	Thr Ser Val Asp Gly Asp Pro His Phe Val Val Asp Phe Pro Leu					
154		680		685		690
155	Ser Arg Leu Thr Val Cys Phe Asn Ile Asp Gly Gln Pro Gly Asp					
156		695		700		705
157	Ile Leu Arg Leu Val Ser Asp His Arg Asp Ser Gly Val Thr Val					
158		710		715		720
159	Asn Gly Glu Leu Ile Gly Ala Pro Ala Pro Pro Asn Gly His Lys					
160		725		730		735
161	Lys Gln Arg Thr Tyr Leu Arg Thr Ile Thr Ile Leu Ile Asn Lys					
162		740		745		750

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/828,423

DATE: 11/26/2001

TIME: 15:04:21

Input Set : N:\Crif3\RULE60\09828423.txt

Output Set: N:\CRF3\11262001\I828423.raw

```

163 Pro Glu Arg Ser Tyr Leu Glu Ile Thr Pro Ser Arg Val Ile Leu
164                               755                               760                               765
165 Asp Gly Gly Asp Arg Leu Val Leu Pro Cys Asn Gln Ser Val Val
166                               770                               775                               780
167 Val Gly Ser Trp Gly Leu Glu Val Ser Val Ser Ala Asn Ala Asn
168                               785                               790                               795
169 Val Thr Val Thr Ile Gln Gly Ser Ile Ala Phe Val Ile Leu Ile
170                               800                               805                               810
171 His Leu Tyr Lys Lys Pro Ala Pro Phe Gln Arg His His Leu Gly
172                               815                               820                               825
173 Phe Tyr Ile Ala Asn Ser Glu Gly Leu Ser Ser Asn Cys His Gly
174                               830                               835                               840
175 Leu Leu Gly Gln Phe Leu Asn Gln Asp Ala Arg Leu Thr Glu Asp
176                               845                               850                               855
177 Pro Ala Gly Pro Ser Gln Asn Leu Thr His Pro Leu Leu Leu Gln
178                               860                               865                               870
179 Val Gly Glu Gly Pro Glu Ala Val Leu Thr Val Lys Gly His Gln
180                               875                               880                               885
181 Val Pro Val Val Trp Lys Gln Arg Lys Ile Tyr Asn Gly Glu Glu
182                               890                               895                               900
183 Gln Ile Asp Cys Trp Phe Ala Arg Asn Asn Ala Ala Lys Leu Ile
184                               905                               910                               915
185 Asp Gly Glu Tyr Lys Asp Tyr Leu Ala Ser His Pro Phe Asp Thr
186                               920                               925                               930
187 Gly Met Thr Leu Gly Arg Gly Met Ser Arg Glu Leu
188                               935                               940

```

193 (2) INFORMATION FOR SEQ ID NO: 2:

195 (i) SEQUENCE CHARACTERISTICS:

196 (A) LENGTH: 3636 base pairs

197 (B) TYPE: nucleic acid

198 (C) STRANDEDNESS: single

199 (D) TOPOLOGY: linear

201 (vii) IMMEDIATE SOURCE:

202 (A) LIBRARY: UTRSNOT02

203 (B) CLONE: 688183

205 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2 :

```

207 CCCTGAGAGC GTCCCGCAGT GGCTGGAGCC CTGGGCGCTG CAAACGTGTC CCGCCGGGTC 60
208 CCCGAGCGTC CCGCGCCCTC GCGCCGCCAT GCTCCTGCTG CTGGGGCTGT GCCTGGGGCT 120
209 GTCCCTGTGT GTGGGGTCGC AGGAAGAGGC GCAGAGCTGG GGCCACTCTT CGGAGCAGGA 180
210 TGGACTCAGG GTCCCGAGGC AAGTCAGACT GTTGCAGAGG CTGAAAACCA AACCTTTGAT 240
211 GACAGAATTC TCAGTGAAGT CTACCATCAT TTCCCGTTAT GCCTTCACTA CGGTTTCCTG 300
212 CAGAATGCTG AACAGAGCTT CTGAAGACCA GGACATTGAG TTCCAGATGC AGATTCCAGC 360
213 TGCAGCTTTC ATCACC AACT TCACTATGCT TATTGGAGAC AAGGTGTATC AGGGCGAAAT 420
214 TACAGAGAGA GAAAAGAAGA GTGGTGATAG GGTAAGAGAG AAAAGGAATA AAACCACAGA 480
215 AGAAAATGGA GAGAAGGGGA CTGAAATATT CAGAGCTTCT GCAGTGATTC CCAGCAAGGA 540
216 CAAAGCCGCC TTTTTCCTGA GTTATGAGGA GCTTCTGCAG AGGCGCCTGG GCAAGTACGA 600
217 GCACAGCATC AGCGTGCGGC CCCAGCAGCT GTCCGGGAGG CTGAGCGTGG ACGTGAATAT 660
218 CCTGGAGAGC GCGGGCATCG CATCCCTGGA GGTGCTGCCG CTTACAACA GCAGGCAGAG 720
219 GGGCAGTGGG CGCGGGGAAG ATGATTCTGG GCCTCCCCCA TCTACTGTCA TTAACCAAAA 780

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/828,423

DATE: 11/26/2001

TIME: 15:04:21

Input Set : N:\Crf3\RULE60\09828423.txt

Output Set: N:\CRF3\11262001\I828423.raw

```

220 TGAAACATTT GCCAACATAA TTTTAAACC TACTGTAGTA CAACAAGCCA GGATTGCCCA 840
221 GAATGGAATT TTGGGAGACT TTATCATTAG ATATGACGTC AATAGAGAAC AGAGCATTGG 900
222 GGACATCCAG GTTCTAAATG GCTATTTTGT GCACTACTTT GCTCCTAAAG ACCTTCCTCC 960
223 TTTACCCAAG AATGTGGTAT TCGTGCTTGA CAGCAGTGCT TCTATGGTGG GAACCAAAC 1020
224 CCGGCAGACC AAGGATGCCC TCTTCACAAT TCTCCATGAC CTCCGACCCC AGGACCGTTT 1080
225 CAGTATCATT GGATTTTCCA ACCGGATCAA AGTATGGAAG GACCACTTGA TATCAGTCAC 1140
226 TCCAGACAGC ATCAGGGATG GGAAAGTGTA CATTCACCAT ATGTCACCCA CTGGAGGCAC 1200
227 AGACATCAAC GGGGCCCTGC AGAGGGCCAT CAGGCTCCTC AACAAGTACG TGGCCACAG 1260
228 TGGCATTTGA GACCGGAGCG TGTCCCTCAT CGTCTTCCTG ACGGATGGGA AGCCCACGGT 1320
229 CGGGGAGACG CACACCTCA AGATCCTCAA CAACACCCGA GAGGCCGCC GAGGCCAAGT 1380
230 CTGCATCTTC ACCATTGGCA TCGGCAACGA CGTGGACTTC AGGCTGCTGG AGAACTGTC 1440
231 GCTGGAGAAC TGTGGCTCA CACGGCGCGT GCACGAGGAG GAGGACGCAG GCTCGCAGCT 1500
232 CATCGGGTTC TACGATGAAA TCAGGACCCC GCTCCTCTCT GACATCCGCA TCGATTATCC 1560
233 CCCCAGCTCA GTGGTGCAGG CCACCAAGAC CCTGTTCCCC AACTACTTCA ACGGCTCGGA 1620
234 GATCATCATT GCGGGGAAGC TGGTGGACAG GAAGCTGGAT CACCTGCACG TGGAGGTCAC 1680
235 CGCCAGCAAC AGTAAGAAAT TCATCATCCT GAAGACAGAT GTGCCTGTGC GGCCTCAGAA 1740
236 GGCAGGGAAA GATGTCACAG GAAGCCCCAG GCCTGGAGGC GATGGAGAGG GGGACACCAA 1800
237 CCACATCGAG CGTCTCTGGA GCTACCTCAC CACAAAGGAG CTGCTGAGCT CCTGGCTGCA 1860
238 AAGTGACGAT GAACCGGAGA AGGAGCGGCT GCGGCAGCGG GCCCAGGCC TGGCTGTGAG 1920
239 CTACCGCTTC CTCCTCCCT TCACCTCCAT GAAGCTGAGG GGGCCGGTCC CACGCATGGA 1980
240 TGGCCTGGAG GAGGCCACG GCATGTCGGC TGCCATGGGA CCCGAACCGG TGGTGCAGAG 2040
241 CGTGCGAGGA GCTGGCACGC AGCCAGGGCC TTTGCTCAAG AAGCCATACC AGCCAAGAAT 2100
242 TAAAATCTCT AAAACATCAG TGGATGGTGA TCCCCACTTT GTTGTGGATT TCCCCCTGAG 2160
243 CAGACTCACC GTGTGCTTCA ACATTGATGG GCAGCCCGGG GACATCCTCA GGCTGGTCTC 2220
244 TGATCACAGG GACTCTGGTG TCACAGTGAA CGGAGAGTTA ATTGGGGCAC CCGCCCCTCC 2280
245 AAATGGCCAC AAGAAACAGC GCACTTACTT GCGCACTATC ACCATCCTCA TCAACAAGCC 2340
246 AGAGAGATCT TATCTCGAGA TCACACCGAG CAGAGTCATC TTGGATGGTG GGGACAGACT 2400
247 GGTGCTCCCC TGCAACCAGA GTGTGGTGGT GGGGAGCTGG GGGCTGGAGG TGTCCGTGTC 2460
248 TGCCAACGCC AATGTCACCG TCACCATCCA GGGCTCCATA GCCTTTGTCA TCCTCATCCA 2520
249 CCTCTACAAA AAGCCGGCGC CCTTCCAGCG ACACCACCTG GGTTTCTACA TTGCCAACAG 2580
250 CGAGGGCCTT TCCAGCAACT GCCACGGAAT GCTGGGTCAG TTCCTGAATC AGGATGCCAG 2640
251 ACTCACAGAA GACCCTGCAG GGCCAGCCA GAACCTCACT CACCCTCTGC TCCTTCAGGT 2700
252 GGGAGAGGGG CCTGAGGCCG TCCTAACAGT GAAAGGCCAC CAAGTCCAG TGGTCTGGAA 2760
253 GCAAAGGAAG ATTTACAACG GGAAGAGCA GATAGACTGC TGGTTTGCCA GGAACAATGC 2820
254 CGCCAAACTG ATTGACGGGG AGTACAAGGA TTACCTGGCA TCCCATCCAT TTGACACAGG 2880
255 GATGACACTT GGCCGGGGAA TGTCCAGGGA GCTCTGAAGC TGGCAGCCTT AAAGATGCAA 2940
256 GTGCATGAAG GACAGTGATG TGGGGAGGCC GTGGGGCAGC TCTTTTCATG GCTTGTACAC 3000
257 GCCTCAGCTC CTGGCAATTA GCTGGACTCC ATGACCCACC CCTGGTGCAG CATAGATCCG 3060
258 ACGTCTGTCT GGGCGAAGGG TAGGGGTGGG TAGGGGCGGG AAGCCTGAGT GCAAATGTCA 3120
259 TTTCCCTCTA CTGCCTCTTC CTGCCTCTCC CCACCTGCC CACATCCACA GAGGGGAGAG 3180
260 AAGGGTCATA GCTAAATGCA ACAAAGTCTG TATCTTGTCC CAACCTGCTT TTCTGTTCTG 3240
261 TTAGCATATC ATAAAGTAAG CCTTCTGGT GAAGGAAGGT TGCTATGAAA CTTTTTTTCT 3300
262 TGGTGGAAAT GGCCAAGTTT AGGCACTCTG CTTTTTGCCT TACACTAATG CTTAGAAAGC 3360
263 TGTCTTTTCA GTGGTGTGTC AGCCCCCAGA TGTGTGGCCA ACCTCTGCTG CAAAGGAATC 3420
264 TCTTGCTGAG TCCAGGCCAC CAATCAGGCA AATAGCCCAT ACATTTGATC GTTGTAACC 3480
265 ATGAAGTCTT TTCTTGCAAG ACGTTTTTCT TCTGCTGTGG TATCTTGCCC TTAAAAATTA 3540
266 GTTTTCATTA AAAAGAAATT TGATTGAAAA TTAATAAAAA ATAAAAAAA AAGAAAAAAA 3600
267 AAAAGAAAGA AAAAATAAAA AAAAAAAA AAAAAA 3636
270 (2) INFORMATION FOR SEQ ID NO: 3:

```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/828,423

DATE: 11/26/2001
TIME: 15:04:22

Input Set : N:\Crf3\RULE60\09828423.txt
Output Set: N:\CRF3\11262001\I828423.raw

L:4 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:10 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

11/26/01